

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR  
SCREENING SIMULATED TRAFFIC FOR RANDOMNESS

Field of the Invention

5 The present invention relates to network traffic simulation and, more particularly, network traffic simulation which includes buy and browse behavior.

Background of the Invention

10 In networked systems, it may be advantageous to simulate network traffic. Conventional network traffic simulation tools exist for providing such simulated traffic. Such simulators are, typically, configured to provide a random stream of network traffic, such as Hypertext Transport Protocol (HTTP) messages which may be bound to a specific network or to other networks.

15 One potential difficulty in network simulation is if the desired simulation has both related and unrelated traffic. For example, in simulating network traffic related to e-commerce, it may be beneficial to simulate buy behavior and browse behavior. Buy behavior is typically characterized as a series of related network traffic messages or packets and browse behavior is typically characterized as a series of unrelated network traffic messages or packets. Typically, conventional network traffic simulators which generate random streams of messages have  
20 difficulty maintaining a random network traffic stream if constrained to provide both buy and browse messages in the network traffic stream.

## Summary of the Invention

Embodiments of the present invention provide methods, systems, and computer program products for simulating network traffic by filtering simulated network traffic utilizing known Ramsey numbers so as to provide a first predefined number of related messages to simulate buy messages and a second predefined number of unrelated messages to simulate browse messages. In particular embodiments of the present invention, the first predefined number corresponds to a clique of order  $m$  specified by a selected known Ramsey number and the second predefined number corresponds to an independent set of order  $n$  specified by the selected known Ramsey number.

In further embodiments of the present invention, simulated network traffic is generated so as to provide a bundle of messages corresponding to the selected known Ramsey number. In such embodiments, filtering the simulated network traffic may be provided by selecting  $m$  related messages from a bundle of messages corresponding to the selected known Ramsey number and by selecting  $n$  unrelated messages from the bundle of messages. Furthermore, the selected  $m$  related messages may be ordered and chained. The  $m$  related messages and the  $n$  unrelated messages may also be streamed so as to provide pseudo-Internet traffic.

In still further embodiments of the present invention, the selection of the  $m$  related and the  $n$  unrelated messages may be monitored to determine if the bundle of messages provides both the  $m$  related messages and the  $n$  unrelated messages. The generation of the bundles of messages may be modified if monitoring of the selection indicates that the bundle of messages fails to provide both the  $m$  related messages and the  $n$  unrelated messages. In such embodiments of the present invention, the generation of simulated network traffic may be provided utilizing a plurality of non-linear functions. In such a case, the generation of the bundle of messages may be modified by adding at least one additional non-linear function to the plurality of non-linear functions.

In still further embodiments of the present invention, the monitoring may be provided by monitoring the selection for a plurality of bundles. The modification of the generation of the network traffic may be provided by modifying

the generation of the bundles of messages if monitoring of the selection indicates that a plurality of the bundle of messages fails to provide both the m related messages and the n unrelated messages.

In still further embodiments of the present invention, the known Ramsey numbers are known Ramsey numbers and/or Ramsey numbers known within bounds.

As will further be appreciated by those of skill in the art, while described above primarily with reference to method aspects, the present invention may be embodied as methods, apparatus/systems and/or computer program products.

10

### Brief Description of the Drawings

**Figure 1** is a block diagram of a data processing system suitable for use in generating simulated network traffic according to embodiments of the present invention;

15 **Figure 2** is a more detailed block diagram of a network traffic simulation system according to embodiments of the present invention;

**Figure 3** is a more detailed block diagram of the network traffic simulation module of **Figure 2**;

20 **Figure 4** is a flowchart illustrating operations for network traffic simulation according to embodiments of the present invention; and

**Figure 5** is a flowchart illustrating operations for network traffic simulation according to further embodiments of the present invention.

### Detailed Description of the Invention

25 The present invention now will be described more fully hereinafter with reference to the accompanying drawings, in which illustrative embodiments of the invention are shown. This invention may, however, be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments are provided so that this disclosure will be thorough and  
30 complete, and will fully convey the scope of the invention to those skilled in the art.

As will be appreciated by one of skill in the art, the present invention may be embodied as a method, data processing system, or computer program product. Accordingly, the present invention may take the form of an entirely hardware embodiment, an entirely software embodiment or an embodiment combining software and hardware aspects all generally referred to herein as a "circuit" or "module." Furthermore, the present invention may take the form of a computer program product on a computer-usable storage medium having computer-usable program code embodied in the medium. Any suitable computer readable medium may be utilized including hard disks, CD-ROMs, optical storage devices, a transmission media such as those supporting the Internet or an intranet, or magnetic storage devices.

Computer program code for carrying out operations of the present invention may be written in an object oriented programming language such as Java®, Smalltalk or C++. However, the computer program code for carrying out operations of the present invention may also be written in conventional procedural programming languages, such as the "C" programming language. The program code may execute entirely on the user's computer, partly on the user's computer, as a stand-alone software package, partly on the user's computer and partly on a remote computer or entirely on the remote computer. In the latter scenario, the remote computer may be connected to the user's computer through a local area network (LAN) or a wide area network (WAN), or the connection may be made to an external computer (for example, through the Internet using an Internet Service Provider).

The present invention is described below with reference to flowchart illustrations and/or block diagrams of methods, apparatus (systems) and computer program products according to embodiments of the invention. It will be understood that each block of the flowchart illustrations and/or block diagrams, and combinations of blocks in the flowchart illustrations and/or block diagrams, can be implemented by computer program instructions. These computer program instructions may be provided to a processor of a general purpose computer, special purpose computer, or other programmable data processing apparatus to produce a

machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks.

5           These computer program instructions may also be stored in a computer-readable memory that can direct a computer or other programmable data processing apparatus to function in a particular manner, such that the instructions stored in the computer-readable memory produce an article of manufacture including instruction means which implement the function/act specified in the flowchart and/or block  
10   diagram block or blocks.

          The computer program instructions may also be loaded onto a computer or other programmable data processing apparatus to cause a series of operational steps to be performed on the computer or other programmable apparatus to produce a computer implemented process such that the instructions which execute on the  
15   computer or other programmable apparatus provide steps for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks.

          Embodiments of the present invention provide for filtering simulated network traffic with known Ramsey numbers so as to simulate buy and browse behavior in the network traffic. Such filtering may potentially provide enhanced  
20   randomness of the simulated network traffic. Simulated network traffic is generated and message bundles having a size equal to the Ramsey number for the desired buy/browse behavior to be simulated are generated. From these message bundles, the buy and browse traffic is selected and streamed to represent pseudo-Internet traffic having the desired buy and browse characteristics.

25           The Ramsey number  $R(m,n)$  gives the solution to the Party Problem, which asks the minimum number of guests  $R(m,n)$  that must be invited so that at least  $m$  will know each other or at least  $n$  will not know each other. Stated differently, the Ramsey number is the minimum number of vertices  $v=R(m,n)$  such that all unidirected simple graphs of order  $v$  contain a clique of order  $m$  and an  
30   independent set of order  $n$ . Ramsey's Theorem states that such a number exists for all  $m$  and  $n$ .

In the present invention, the clique of order m is considered the buy traffic and the independent set of order n the browse traffic. Some known Ramsey numbers are illustrated in Table 1 below.

5 Table 1. Exemplary Known Ramsey Numbers

| m | n | R(m,n) |
|---|---|--------|
| 3 | 3 | 6      |
| 3 | 4 | 9      |
| 3 | 5 | 14     |
| 3 | 6 | 18     |
| 3 | 7 | 23     |
| 3 | 9 | 36     |
| 4 | 4 | 18     |
| 4 | 5 | 25     |

15 Additional information on Ramsey numbers, including a summary of known Ramsey numbers, may be found, for example, in Radziszowski *et al.*, "Small Ramsey Numbers", Revision # 7, The Electronic Journal of Combinatorics (2000).

Various embodiments of the present invention will now be described with reference to the figures. **Figure 1** illustrates an exemplary embodiment of a data processing system **130** suitable for simulating network traffic in accordance with embodiments of the present invention. The data processing system **130** typically includes input device(s) **132** such as a keyboard or keypad, a display **134**, and a memory **136** that communicate with a processor **138**. The data processing system **130** may further include a speaker **144**, and an I/O data port(s) **146** that also communicate with the processor **138**. The I/O data ports **146** can be used to transfer information between the data processing system **130** and another computer system or a network. These components may be conventional components, such as those used in many conventional data processing systems, which may be configured to operate as described herein.

**Figure 2** is a block diagram of data processing systems that illustrates systems, methods, and computer program products in accordance with embodiments of the present invention. The processor **138** communicates with the memory **136** via an address/data bus **248**. The processor **138** can be any commercially available or custom microprocessor. The memory **136** is representative of the overall hierarchy of memory devices containing the software and data used to implement the functionality of the data processing system **130**. The memory **136** can include, but is not limited to, the following types of devices: cache, ROM, PROM, EPROM, EEPROM, flash memory, SRAM, and DRAM.

As shown in **Figure 2**, the memory **136** may include several categories of software and data used in the data processing system **130**: the operating system **252**; the application programs **254**; the input/output (I/O) device drivers **258**; and the data **256**. As will be appreciated by those of skill in the art, the operating system **252** may be any operating system suitable for use with a data processing system, such as OS/2, AIX or System390 from International Business Machines Corporation, Armonk, NY, Windows95, Windows98 or Windows2000 from Microsoft Corporation, Redmond, WA, Unix or Linux. The I/O device drivers **258** typically include software routines accessed through the operating system **252** by the application programs **254** to communicate with devices such as the I/O data port(s) **146**, the data storage **135** and certain memory **136** components. The application programs **254** are illustrative of the programs that implement the various features of the data processing system **130** and preferably include at least one application which supports operations according to embodiments of the present invention. Finally, the data **256** represents the static and dynamic data used by the application programs **254**, the operating system **252**, the I/O device drivers **258**, and other software programs that may reside in the memory **136**.

As is further seen in **Figure 2**, the application programs **254** may include an Internet traffic simulation module **260**. The Internet traffic simulation module **260** may carry out the operations described herein for generating simulated Internet traffic which includes both buying and browsing behavior. The data portion **256** of memory **136**, as shown in the embodiments of **Figure 2**, may include known

Ramsey numbers **262**. The known Ramsey numbers **262** may be utilized by the Internet traffic simulation module **260** to filter stochastic and/or chaotic simulated Internet traffic to provide substantially random Internet traffic which includes buy and browse behavior.

5 While the present invention is illustrated, for example, with reference to the Internet traffic simulation module **260** being an application program in **Figure 2**, as will be appreciated by those of skill in the art, other configurations may also be utilized while still benefitting from the teachings of the present invention. For example, the Internet traffic simulation module **260** may also be incorporated into  
10 the operating system **252**, the I/O device drivers **258** or other such logical division of the data processing system **130**. Thus, the present invention should not be construed as limited to the configuration of **Figure 2** but is intended to encompass any configuration capable of carrying out the operations described herein.

**Figure 3** is a more detailed block diagram of the traffic simulation module  
15 **260** of **Figure 2**. As seen in **Figure 3**, a message traffic generation unit **300** generates message bundles which are provided to a Ramsey function filter **310**. The message bundles preferably include  $R(m,n)$  messages, however, a stream of messages may be provided and the Ramsey function filter **310** may divide the stream into bundles of  $R(m,n)$  messages. The message traffic generation unit **300**  
20 may be a set of non-linear functions producing stochastic and/or "chaotic" message traffic. Techniques for generating network traffic will be understood by those of skill in the art in light of the present disclosure and, therefore, will not be described further herein except as they relate to particular aspects of the present invention. The message traffic generated by the traffic generation unit **300** is of a type from  
25 within the realm of buy/browse behaviors.

The Ramsey function filter **310** receives the message bundles or message stream from the message traffic generation unit **300** and identifies a clique of order  $m$  and an independent set of order  $n$ , where  $m$  is the number of related e-commerce buy messages needed to complete a particular buy behavior and  $n$  is the number of  
30 browse messages. The remaining messages in the bundle of messages from the message traffic generation unit **300** may be discarded by the Ramsey function filter



310. The values for  $m$  and  $n$  may be selected based on a desired and/or expected buy/browse behavior. Selecting the desired buy/browse behavior to be simulated, thus, selects the Ramsey number  $R(m,n)$ .

Optionally, the Ramsey function filter 310 may provide feedback to the message traffic generation unit 300 so as to adjust the functions of the message traffic generation unit 300 if, for example, the Ramsey function filter 310 is not configured to generate the desired messages from the message bundles provided.

The Ramsey function filter 310 provides the  $m$  related messages and the  $n$  unrelated messages to the streaming module 320. The streaming module 320 chains and orders the related messages to simulate the buy behavior and streams the related and unrelated messages to represent pseudo-Internet traffic. Such traffic may, for example, be provided to a border router of an e-commerce system so as to simulate buy and browse behavior traffic to the e-commerce system.

Operations according to embodiments of the present invention will now be described with reference to the flowcharts of **Figures 4 and 5**. As seen in **Figure 4**, stochastic and/or chaotic message traffic is generated (block 400), for example, by the message traffic generation module 300. The generated traffic is filtered utilizing known Ramsey numbers to select from the generated traffic  $m$  related messages and  $n$  unrelated messages (block 402). For example, message bundles of size  $R(m,n)$  can be generated and evaluated so as to select the  $m$  related messages and the  $n$  unrelated messages. The remaining messages could be discarded. In any event, the  $m$  related messages are chained and ordered to simulate the buy messages (block 404) and streamed with the  $n$  unrelated messages to represent pseudo-Internet traffic (block 406).

**Figure 5** illustrates further embodiments of the present invention which provide feedback to the message traffic generation unit 300 so as to facilitate Ramsey filtering of the generated message traffic. As with the embodiments illustrated in **Figure 4**, stochastic and/or chaotic message traffic is generated (block 500), for example, by the message traffic generation module 300. The generated traffic is filtered utilizing known Ramsey numbers to select from the generated traffic  $m$  related messages and  $n$  unrelated messages (block 502). For example,

message bundles of size  $R(m,n)$  can be generated and evaluated so as to select the  $m$  related messages and the  $n$  unrelated messages. The remaining messages could be discarded.

The operation of the Ramsey function filter **310** may be monitored to  
5 determine if the messages are following the Ramsey function (e.g., are  $m$  related and  $n$  unrelated messages present in the  $R(m,n)$  bundles of messages) (block **504**). If the messages are not satisfying the Ramsey function (block **504**), the message traffic generation unit **300** may be modified, for example, by adding one or more additional non-linear functions to the non-linear functions utilized to generate the  
10 message traffic (block **506**). Additional message traffic is then generated utilizing the modified non-linear functions (block **500**). Alternatively, the non-linear functions utilized in message generation could be modifying by revising the existing non-linear functions or replacing the existing non-linear functions utilized with other non-linear functions.

15 Such monitoring and evaluation of the Ramsey function filtering and feedback for message generation may be performed on each individual message bundle or on a plurality of message bundles. If a single bundle of  $R(m,n)$  does not provide messages which satisfy the Ramsey function, multiple bundles could be evaluated to select the  $m$  buy messages and  $n$  browse messages. The feedback  
20 would then, over time, adjust the generation of messages so as to provide messages such that all or substantially all of the message bundles satisfy the Ramsey function.

When messages which satisfy the Ramsey function are available (block **504**), the  $m$  related messages are chained and ordered to simulate the buy messages  
25 (block **508**) and streamed with the  $n$  unrelated messages to represent pseudo-Internet traffic (block **510**).

While the present invention has been described with reference to the use of known Ramsey numbers, as used herein, references to known Ramsey numbers refers to both known Ramsey numbers and Ramsey numbers known within  
30 bounds.

0001107-073304

The flowcharts and block diagrams of **Figures 1** through **6** illustrate the architecture, functionality, and operation of possible implementations of systems, methods and computer program products for generating simulated Internet traffic according to various embodiments of the present invention. In this regard, each  
5 block in the flow charts or block diagrams may represent a module, segment, or portion of code, which comprises one or more executable instructions for implementing the specified logical function(s). It should also be noted that, in some alternative implementations, the functions noted in the blocks may occur out of the order noted in the figures. For example, two blocks shown in succession  
10 may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, depending upon the functionality involved. It will also be understood that each block of the block diagrams and/or flowchart illustrations, and combinations of blocks in the block diagrams and/or flowchart illustrations, can be implemented by special purpose hardware-based systems  
15 which perform the specified functions or acts, or combinations of special purpose hardware and computer instructions.

In the drawings and specification, there have been disclosed typical illustrative embodiments of the invention and, although specific terms are employed, they are used in a generic and descriptive sense only and not for  
20 purposes of limitation, the scope of the invention being set forth in the following claims.